venio systems

Defensibility: The Secret Facts of eDiscovery

Prelude

You've just gotten that dreaded call from the eDiscovery Administrator – It's going to take another two days to get the data up for review. They mumble something incomprehensible about "container extraction, corrupt signatures". After you hang up the phone, questions linger. It's 10 o'clock – Do you know where your eDiscovery is?

The Challenge

In a litigation, the producing party and their attorney of record are ultimately responsible for handling eDiscovery. But because the traditional eDiscovery solutions were nothing more than a complex set of legacy applications cobbled together, the results are error prone manual tasks, inaccurate data culling, and inefficient review cycles that lead to high cost and low productivity of most expensive resources.

The eDiscovery industry has grown up supported by legacy applications that require enormous amounts of technical know-how and project manager overview. Most organizations do not use end-to-end processing and review tools, instead, they use multiple eDiscovery software applications that require manual tasks and the duplication of data.

Because eDiscovery processing issues don't get enough attention in the press, these are called "The Secret Facts of eDiscovery". These problems are creating data spoliation between collection and review and create a risk to the chain of custody.

eDiscovery Processing

Secret Fact #1: Processing Extraction Levels

eDiscovery processing is a set of tasks that software performs to take collected data such as, a Microsoft email (.pst) file and extracts the individual components. Each component is referred to as an "item" and each "item" must be able to be searched, filtered, reviewed, and/or produced during eDiscovery. Unfortunately, items contain their own components such as attachments, nested attachments, zipped or embedded files. It is easy for a single email to extract out to numerous items at multiple levels.

FLAWED SEARCH AND PRIVILEGE REVIEW

If the processing software has a level limit, it may not index all the attachments or embedded files. Searches used to find potentially responsive data will miss these non-indexed files, and they will not go on to review unless they are part of items that are returned by a search term.
During review, non-extracted embedded files may not be reviewed at all, if a reviewer doesn't open the embedded document and look at it in its entirety. The reviewer may not see potentially privileged, non-responsive or harmful content, and those embedded files may be produced without any review.

The Solution to Secret #1

Ask the technologist or provider if there are level limits on any of the applications being used in the eDiscovery workflow. Do not be satisfied with a limit that sounds like some huge number or the answer, "We've never had a problem."

The only reliable answer is "none".

Multiple Applications

Secret Fact #2: Multiple Processing Applications

Many processing operations have been built on outdated workflows that require a whole tool chest of eDiscovery applications. Whenever more than one tool is used the data is exported from (Tool A) into a file and that new file is imported into (Tool B). Potential problems emerge any time eDiscovery is exported and reimported. In one Current Model used in many processing environments, the data or a portion of the data, can be exported and imported multiple times before it is produced.

Why is this so common? Because some older applications cannot perform all eDiscovery tasks accurately or cost effectively, and processors use cheaper tools early in the project where the data volume is higher.

Data Loss:

As with many of these "Secret Facts", human interaction in processing is risky. When exporting from Tool A to Tool B, an eDiscovery Administrator may need to kick-off the process and make sure that the export is complete and that the file is imported accurately. This may sound simple, but is a breeding-ground for spoliation. Tool A and Tool B may not have their settings aligned, and metadata such as dates may be modified or dropped.

Organizations may advertise that their processing environment is integrated, yet still use multiple tools. What really happens is that "middleware" is performing the data transfer from Tool A to Tool B. Providers may even say "everything's in the cloud" as a way of implying that the processing method is fully integrated. If more than one application is being used in the cloud, it's still not integrated.

The Solution to Secret Fact #2: Technology Selection

The savvy litigator or eDiscovery professional needs to look under the hood of any processing method and understand just how many applications or modules are deployed that require data transfers. Selection of vendors or solutions should include a comparison of end-to-end solutions that perform all processing and review tasks without moving or modifying data to pseudo-integrated offerings that require automated export/import tasks. The same buyer rules apply, related to chain-of-custody and defensibility when working in the cloud. Consumers need to understand all the components of any "integrated" or cloud offering that may not be built on an end-to-end platform.

Multiple Hardware Platforms

Secret Fact #3: Multiple Hardware Platforms

The story of the phone call from the prelude really happened. The real reason for the two-day delay was that a server crashed when moving data. The problem with multiple applications was exacerbated because they were on different hardware platforms hosted by the same vendor. There were hundreds of gigabytes of data moving from Server A on the Tenth Floor to Server B on the Eleventh, and when Server A crashed, the whole thing had to be restarted. This generated concerns that data would be unrecoverable.

Time and Money

Moving data takes time. Moving data between servers can slow down the team's ability to meet deadlines and add significantly to labor costs. Moving data from one cloud instance to another may be faster than from one physical server to another, but it will still require wasted processing time.

The Solution to Secret Fact #3

In regards to techology selection and hardware planning, eDiscovery professionals need to scrutinize the hardware requirements of a platform for multiple applications. An eDiscovery process that takes multiple hardware installations, requires multiples of everything: expensive expertise, costly maintenance, and their related disaster recovery infrastructure. Users of cloud services must evaluate the various methods of getting the data into the cloud. A truly efficient cloud application will have a process for direct data transfer without an intermediate step of transfer to a cloud server and then ingesting into the processing application.

Human Partcipation

Secret Fact #4: Human Participation

There is a huge amount of human interaction in the eDiscovery process, and this is the source of most of the problems that occur during processing and review:

CONTINUITY:

A common failure is the lack of standardization from one batch of ESI to another on the same project. Data from different sources can be loaded over months or years. Without stored templates or built-in workflow the batches may be processed differently which can cause many of the chain of custody and data problems discussed earlier. Project teams may also have to remember to search for items that need OCR, tiffing or slip-sheets and to make sure and run the same responsive or privileged search terms on each batch.

METRICS:

Many eDiscovery tools are just not “user friendly” because there were designed to be used solely by technicians and may have a limited list of canned reports that are difficult and time-consuming run. It is no wonder that litigators have problems tracking time and costs when metrics are hard to generate. This compromises the ability of the project manager or litigator to understand the data and supervise the work.

The Solution to Secret Fact #4

AUTOMATED WORKFLOW:

Applications should include templates that save settings and perform tasks proactively such as automatically OCRing necessary files and indexing them. An automated workflow can repeat settings, correct errors and perform tasks for every batch of a project resulting in a process which is easier to document, defend, and communicate.

REAL-TIME REPORTING: DASHBOARDS

Good reports, easily generated, are essential. But why wait? Why not just keep an eye on the project through a dashboard showing real-time stats through easily understandable visual displays? Dashboard functionality, such as the Venio examples below, is becoming the standard in team communication.

REAL-TIME REPORTING: DASHBOARDS

Good reports, easily generated, are essential. But why wait? Why not just keep an eye on the project through a dashboard showing real-time stats through easily understandable visuals? Dashboard functionality, such as the Venio example below, is becoming the standard in team communication.

In addition, some applications port metrics to mobile devices.

STATUS REPORTS:

Reports are still invaluable even with dashboards. A good eDiscovery solution will have a full suite of helpful reports to run with minimal tech interaction. eDiscovery selection teams must see all user-generated reports, including Exception Reports to make sure that important information can be generated directly by the user.

Keyword Searching

Secret Fact #5: Keyword Searching

Predictive coding may have gained ground as a method for limiting the volume of human review, but many projects still include keyword searching to find responsive items, likely non-responsive items, and potentially privileged documents. Many teams perform these operations in an ad hoc, haphazard manner risking the completeness of the production.

These teams have still not benefited from Magistrate Judge Paul Grimm's search "how to" outlined in Victor Stanley, Inc. v. Creative Pipe, Inc., 250 F.R.D. 251 (D. Md. 2008). Parties either botch the search process so as to make it indefensible or agree to overly broad keywords during the meet and confer process.

SEARCH APPLICATIONS

Magistrate Judge John M. Facciola cautioned in U.S. v. Michael John O'Keefe, 537 F. supp. 2d 14 (2008) that "this topic is clearly beyond the ken of a layman" advising litigants to get expert advice when dealing with eDiscovery issues such as searching.

MULTIPLE APPLICATIONS

When performing keyword searches in two applications, the eDiscovery Specialist must translate the original search terms from Tool A into the correct syntax for Tool B. This is hampered by the differing search methods of each search engine. If the two applications are being used for finding potentially responsive items, the number of responsive items to the patent number search will not match.

METADATA SEARCHES

Inexperienced searchers may forget to search the metadata and assume that the keyword will run on the contents of an item as well as the metadata fields. This is true of some search software but not all. Users need to be clear as to exactly which fields or parts of an item are being searched.

LEVEL LIMITS

It must be emphasized that accurate search can only be performed if all items, including attachments and embedded files, are fully indexed. Email archives and/or document management systems that don't enable searching of every piece of stored information or eDiscovery applications, that cannot go beyond a certain level when indexing files, are not viable search tools.

REPORTING AND DOCUMENTATION

Many search methods are not designed to defend the process through detailed reports that track every iteration of each search term. The search may miss items due to misspellings or otherwise incorrect terms or include overly broad terms which will produce unrelated documents resulting in data sprawl.

The Solution to Secret #5

SEARCH EXPERTISE

Take Judge Facciola's advice, if there is insufficient in-house expertise, hire an expert such as a Data Analytics Consultant. Their work will pay for itself in reduced review volumes and produce a higher level of comfort with defensibility and privilege.

SEARCH ENGINE

Make sure and understand if more than one search engine will be used during the project and ask how they differ. Also ask as to how the search results and their reports will be codified to show clear chain of custody throughout the process. Consider an end-to-end solution that filters and searches data from ingestion through to production to enable consistent searching and consistent results.

REPORTS AND DOCUMENTATION

Review report formats for keyword searching and documentation of the search method when selecting a vendor, or at the start of the project to ensure that the software can generate detailed reports. In the example Venio search report below, several important metrics are delivered for each individual term.
Documentation isn't just a luxury; it may be the only thing supporting non-waiver of privilege. In Clark Cnty. v Jacobs Facilities, Inc., No. 2:10-cv-00194-LRH-PAL, 2012 WL 4609427 (D. Nev. Oct. 1, 2012), privilege was not waived because the responding party was able to document their search method showing each individual search term.

Conclusion

Your responsibility

These may be "Secret Facts" to some but they undermine a responsibility that cannot be shirked. The eDiscovery industry is constantly raising the level of technology that supports and automates best practices through end-to-end workflow and the legal community and courts have put a great deal of effort into amending standards to reflect the importance of ESI and its correct management in Twenty-first century discovery practice.

About Venio Systems

At Venio Systems, we are dedicated to working with our trusted partners to bring the latest legal technology innovations to law firms, agencies, and corporations. We provide e-Discovery solutions that streamline and improve the litigation process for our clients and partners, allowing them to achieve the most successful outcome possible while saving time and resources.

Experience the VenioOne difference with our unified platform that powers every phase of e-Discovery: processing, early case assessment, legal hold, legal analysis, culling, review, and digital production. Our VenioOne OnDemand self-service solution provides users an agile and easy to use system throughout the litigation process. Our Venio Cloud product offers a turn-key, full featured solution which can handle cases of any size or complexity.

Ask how Venio Systems can handle all of your eDiscovery needs from ingestion to production, all under one platform. Contact us today to see a no obligation demonstration.